STUDIES ON INBREEDING. VII.—SOME FURTHER CONSIDERATIONS REGARDNG THE MEASUREMENT AND NUMERICAL EXPRESSION OF DEGREES OF KINSHIP[1]

DR. RAYMOND PEARL

1. IN this series of studies certain concepts regarding the quantitative aspect of inbreeding have been presented. These concepts have in part been rigorously defined, and expressed in mathematical form. It is desirable to repeat here and extend in certain directions, the definition of two of the most fundamental of these concepts.

I. *Inbreeding* is defined in these studies as the condition or state in which an organism has in fact fewer different ancestors than the maximum number possible.

The degree or amount of inbreeding (total) is measured by a series of *inbreeding coefficients,* one for each ancestral generation, defined by the following equation:

$$Z_n = \frac{100\ (p_{n+1} - q_{n+1})}{p_{n+1}}, \tag{i}$$

where p_{n+1} denotes the maximum possible number of different individuals involved in the matings of the $n+1$ generation, q_{n+1} the *actual* number of different individuals involved in these matings, and Z_n is the inbreeding coefficient for the $n+1$-th ancestral generation.

II. A state or condition of *relationship* or kinship between two organisms exists when these organisms have one or more common ancestors. The degree, intensity or closeness of the relationship is, in general, proportional to the number of different ancestors which the two individuals have in common, out of the whole number they might possibly have in common.

[1] Papers from the Biological Laboratory of the Maine Agricultural Experiment Station. No. 113.

The degree or amount of relationship, in accordance with the above definition, is numerically measured by *relationship coefficients,* one for each ancestral generation. The coefficients are calculated in two slightly different ways according to whether they are being evaluated in connection with inbreeding coefficients, which will usually be the case, or independently.

A. When calculated in connection with inbreeding coefficients, a relationship coefficient is calculated, by methods presently to be shown by example, in accordance with the following equation:

$$\frac{K_n}{100} = \frac{(p_{n+1} - q_{n+1}) - ({}_sZ_{n-1} \cdot {}_sp_n + {}_dZ_{n-1} \cdot {}_dp_n)}{\frac{1}{2}p_{n+1}}, \qquad \text{(ii)}$$

where the letters have the same significance as in (i) with the additions that K denotes a relationship coefficient, a prefixed subscript s means that letters following it refer to the pedigree of the sire only, and a prefixed subscript d means that the letters following refer to the pedigree of the dam only.

B. When calculated independently of inbreeding coefficients, as, for example, to measure the relationship between two male animals, the relationship coefficient becomes

$$\frac{K_n}{100} = \frac{p_{n+1} - r_{n+1}}{\frac{1}{2}p_{n+1}}, \qquad \text{(iii)}$$

where $p_{n+1} - r_{n+1}$ denotes the number of ancestors in the $n+1$-th generation (each individual and its ancestry being counted once only) which occur, in the $n+1$-th or some earlier ancestral generation, in the pedigrees of both animals, or in other words which are common ancestors; p_{n+1} denotes the *total* number of ancestors in the same generation of both pedigrees taken together.

III. Inbreeding, defined in I, may exist in respect of any individual, as a result of any one or a combination of the following circumstances: (a) the sire of the individual has fewer than the maximum possible number of different ancestors, and no ancestors in common with the dam; (b)

the dam of the individual has fewer than the maximum possible number of different ancestors, and no ancestors in common with the sire; or, (c) the sire and dam have a certain number of common ancestors, and hence are, in the common sense of the word, related to each other in some degree.

IV. We may separate conceptually that portion of the total inbreeding due to a or b or any combination of a and b, from that portion of the total inbreeding due to c, and define as due to *relationship* between the sire and dam that amount or degree of inbreeding (in the sense of I) which remains after the amount due to a or b (of III) or any combination of a and b has been subtracted from the total inbreeding.

A numerical expression of the portion of the inbreeding in the nth generation due to relationship is obtained by a *partial inbreeding index* of the following form:

$$KZ_n = \frac{50(K_n)}{Z_n}. \tag{iv}$$

Expressed in words this means that we take as an index of the part of the inbreeding due to relationship the percentage which one half of the relationship coefficient is of the inbreeding coefficient, both referred of course to the same ancestral generation.

2. The above paragraphs define a relationship coefficient much more rigorously and generally than was done in my earlier paper on the subject,[2] or in "Modes of Research in Genetics."[3] Not only is this a gain in itself, but also it makes possible a great simplification in the actual work of calculating coefficients of relationship from pedigrees. Extensive experience has shown that the method of making these determinations given in my earlier paper left much to be desired in the direction of simplicity, ease of application, and even of accuracy in case the pedigree dealt with was at all complicated in

[2] Pearl, R., AMER. NAT., Vol. XLVIII, pp. 513–523, 1914.

[3] Pearl, R., "Modes of Research in Genetics," New York, 1915 (Macmillan & Co.). Cf. pp. 101–156.

respect of the distribution of its ancestral repetition. Out of actual laboratory experience has been developed the more simple and rigorous analysis of the matter presented in this paper.

3. It would appear that the briefest and simplest way to make clear our concept of kinship measurements, its use in the analysis of inbreeding, and its practical application to pedigrees, is to carry out the work on some concrete examples, given by actual pedigrees showing a rather high degree of inbreeding or relationship. This we shall accordingly proceed at once to do, taking as our first example the pedigree through five ancestral generations of the Jersey cow Letty's Fancy Lady (241551).

The pedigree (for five ancestral generations) of this cow is presented in Tables I and II. Table I gives the pedigree of her *sire*, Rioter's St. Lambert King (58644), and Table II gives the pedigree of the *dam* of the cow, Letty's Fancy (160320). Tables I and II together, therefore, give the complete pedigree (to the extent already indicated) of the cow herself. The reason for splitting the pedigree into two parts in this way in its presentation will be apparent as we proceed. The numbers preceding the names of the animals are the registry numbers in the Herd Books of the American Jersey Cattle Club.

In Tables I and II the symbols have the following significance: A solid circle indicates a primary reappearance of an ancestor, having reference to the pedigree of Letty's Fancy Lady *as a whole,* and an open circle indicates an entailed reappearance consequent upon the primary reappearance denoted by the solid circle. A solid square indicates a primary reappearance in the pedigree of the *sire* of Letty's Fancy Lady, *considered by itself and without reference to her dam's pedigree;* an open square denotes reappearance consequent upon those indicated by the solid squares. Finally, a solid diamond indicates a primary reappearance of an ancestor in the pedigree of the *dam* of Letty's Fancy Lady, *considered by itself,* while the open diamonds denote the corresponding entailed reappearances.

TABLE I

PEDIGREE OF RIOTER'S ST. LAMBERT KING (58644), SIRE OF LETTY'S FANCY LADY (241551)

Ancestral Generation [4]				
1	2	3	4	5
Sex ♂ No. 58644 Rioter's St. Lambert King	♂ No. 16501 Rioter of St. Lambert	No. 15175 ♂ King of St. Lambert	No. 13656 ♂ Ida's Rioter of St. Lambert	No. 4558 ♂ ● Bachelor of St. Lambert
				No. 24990 ♀ Ida of St. Lambert
			No. 24991 ♀ Allie of St. Lambert	No. 2238 ♂ ●■ Stoke Pogis 3d
				No. 5122 ♀ Kathleen of St. Lambert
		No. 28353 ♀ May Day Stoke Pogis	No. 2238 ♂ Stoke Pogis 3d	No. 1259 ♂ Stoke Pogis
				No. 3239 ♀ Marjoram
			No. 5109 ♀ May Day of St. Lambert	No. 1066 ♂ Lord Lisgar
				No. 1373 ♀ Jerne
	♀ No. 78867 Phyllis of St. Lambert	No. 15175 ♂ ●■ King of St. Lambert	No. 13656 ♂ ○□ Ida's Rioter of St. Lambert	No. 4558 ♂ ○□ Bachelor of St. Lambert
				No. 24990 ♀ ○□ Ida of St. Lambert
			No. 24991 ♀ ○□ Allie of St. Lambert	No. 2238 ♂ ○□ Stoke Pogis 3d
				No. 5122 ♀ ○□ Kathleen of St. Lambert
		No. 43671 ♀ Allie of St. Lambert 2d	No. 8388 ♂ Canada's John Bull	No. 6036 ♂ Sir George of St. Lambert
				No. 12968 ♀ Nymph of St. Lambert
			No. 24991 ♀ ●■ Allie of St. Lambert	No. 2238 ♂ ○□ Stoke Pogis 3d
				No. 5122 ♀ ○□ Kathleen of St. Lambert.

[4] Referred to the propositus, Letty's Fancy Lady (241551).

TABLE II

PEDIGREE OF LETTY'S FANCY (160320), DAM OF LETTY'S FANCY LADY (241551)

Ancestral Generations				
1	2	3	4	5
Sex ♀ Letty's Fancy No. 160320	♂ No. 48228 Rioter's Exile of St. Lambert	No. 13657 ♂ Exile of St. Lambert	No. 4558 ♂ Bachelor of St. Lambert	No. 3143 ♂ Orloff
				No. 6638 ♀ Charity of St. Lambert
			No. 24991 ♀ ● Allie of St. Lambert	No. 2338 ♂ ○ Stoke Pogis 3d
				No. 5122 ♀ ○ Kathleen of St. Lambert
		No. 73475 ♀ Letty Rioter	No. 10481 ♂ Diana's Rioter	No. 6036 ♂ ● Sir George of St. Lambert
				No. 6636 ♀ Diana of St. Lambert
			No. 48128 ♀ Letty Coles 2d	No. 10481 ♂ ● ◆ Diana's Rioter
				No. 23351 ♀ Letty Coles
	♀ No. 142135 Lady Letty Rioter	No. 17408 ♂ St. Lambert Boy	No. 8388 ♂ ● Canada's John Bull	No. 6036 ♂ ○ Sir George of St. Lambert
				No. 12968 ♀ ○ Nymph of St. Lambert
			No. 14880 ♀ Oakland's Nora	No. 5248 ♂ Lorne
				No. 5123 ♀ Pet of St. Lambert
		No. 124201 ♀ Lady Letty Lambert	No. 17408 ♂ ● ◆ St. Lambert Boy	No. 8388 ♂ ○ ◇ Canada's John Bull
				No. 14880 ♀ ○ ◇ Oakland's Nora
			No. 48128 ♀ ● ◆ Letty Coles 2d	No. 10481 ♂ ○ ◇ Diana's Rioter
				No. 23351 ♀ ○ ◇ Letty Coles

With these data in hand we may proceed to the evaluation first of the total inbreeding. We have in Table III the pedigree elimination table for this purpose, which lists the primary reappearances indicated by solid circles.

TABLE III

PEDIGREE ELIMINATION TABLE FOR THE TOTAL INBREEDING OF LETTY'S FANCY LADY

Name of Animal Primarily Reappearing	Ancestral Generation in which Primary Reappearance Occurs				
	1	2	3	4	5
King of St. Lambert	—	—	1	2	4
Allie of St. Lambert	—	—	—	2	4
Canada's John Bull	—	—	—	1	2
St. Lambert Boy	—	—	—	1	2
Letty Coles 2d	—	—	—	1	2
Bachelor of St. Lambert	—	—	—	—	1
Stoke Pogis 3d	—	—	—	—	1
Sir George of St. Lambert	—	—	—	—	1
Diana's Rioter	—	—	—	—	1
Totals	0	0	1	7	18

Whence, by the usual method, using the tables of Pearl and Miner,[5] we have the following values:

TOTAL INBREEDING COEFFICIENTS FOR LETTY'S FANCY LADY

$$Z_1 = 0, \quad Z_2 = 12.50, \quad Z_3 = 43.75, \quad Z_4 = 56.25.$$

Let us next consider Table IV, which gives the pedigree elimination for the pedigree of the sire, as given in Table I, considered by itself, the primary reappearances listed being those indicated by solid squares. It must be particularly noted that the primary reappearances listed in this table are referred to the ancestral generations of the pedigree of Letty's Fancy Lady, and *not* to the pedigree of Rioter's St. Lambert King, her sire, with whose pedigree we are dealing.

TABLE IV

PEDIGREE ELIMINATION TABLE FOR RIOTER'S ST. LAMBERT KING

Name of Animal Primarily Reappearing	Ancestral Generation[6] in which Primary Reappearance Occurs				
	1	2	3	4	5
King of St. Lambert	—	—	1	2	4
Allie of St. Lambert	—	—	—	1	2
Stoke Pogis 3d	—	—	—	—	1
Totals	0	0	1	3	7

[5] Pearl, R., and Miner, J. R., Maine Agr. Expt. Stat. Ann. Rept. for 1913, pp. 191–202.

[6] Referred to the pedigree of Letty's Fancy Lady.

In Table V exactly corresponding data are given for the pedigree of Letty's Fancy, the dam of Letty's Fancy Lady. The primary reappearances here are those indicated by solid diamonds in Table II.

TABLE V

PEDIGREE ELIMINATION TABLE FOR LETTY'S FANCY

Name of Animal Primarily Reappearing	Ancestral Generation[6] in which Primary Reappearance Occurs				
	1	2	3	4	5
St. Lambert Boy	—	—	—	1	2
Letty Coles 2d	—	—	—	1	2
Diana's Rioter	—	—	—	—	1
Totals	0	0	0	2	5
Combined Totals of Tables IV and V	0	0	1	5	12
Difference between combined totals, and totals of Table III (total inbreeding)	0	0	0	2	6

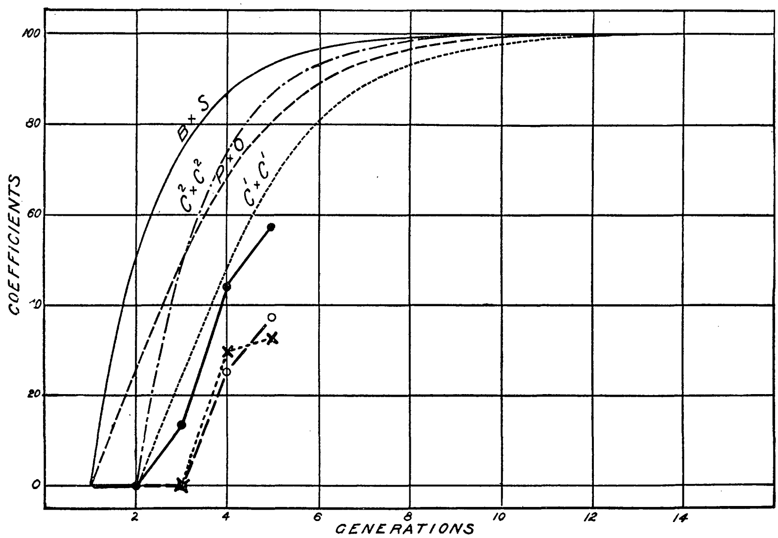

FIG. 1. Diagram showing the inbreeding and relationship curves for Letty's Fancy Lady. Total inbreeding coefficients—solid line and circles; relationship coefficients—dash line and open circles; partial inbreeding coefficients—dotted line and crosses. The smooth curves at the top of the diagram are the total inbreeding curves for continued brother × sister, parent × offspring, and single and double cousin × cousin mating. These are inserted for comparison.

From the last line of Table V we deduce *relationship coefficients* as follows:

$$K_1 = 0, \qquad K_3 = \frac{100 \times 2}{8} = 25.00,$$

$$K_2 = 0, \qquad K_4 = \frac{100 \times 6}{16} = 37.50.$$

Expressed in words these coefficients mean that Rioter's St. Lambert King and Letty's Lady are related to the amount or degree of 25 per cent. in the third, and 37.5 per cent. in the fourth ancestral generation.

In Fig. 1 are shown the total inbreeding (solid line and dots), the relationship (dashes and open circles), and the partial inbreeding (dots and crosses) curves for Letty's Fancy Lady.

Finally we have, from (iv), the following coefficients of partial inbreeding due to relationship.

$$KZ_1 = \frac{50\ (0)}{0} = 0,$$

$$KZ_2 = \frac{50\ (0)}{12.5} = 0,$$

$$KZ_3 = \frac{50\ (25)}{43.75} = 28.57,$$

$$KZ_4 = \frac{50\ (37.5)}{56.25} = 33.33.$$

We thus see that of the total inbreeding observed in the third ancestral generation of Letty's Fancy Lady, none is due to relationship between her sire and dam; of that observed in the fourth ancestral generation, 28.57 is due to such relationship; and finally, of that observed in the fifth ancestral generation, one third arises because of relationship between sire and dam.

4. Let us next consider an example of measuring relationship independently, altogether apart from consideration of inbreeding. We may take a very simple case afforded by the two milking shorthorn cows, Imp. Milk

Maid 211032, and Imp. White Queen 545726. The pedigrees of these animals follow in Tables VI and VII. The problem before us is to measure and express numerically the degree of relationship or kinship between these two animals.

TABLE VI

Pedigree of Imp. Milk Maid (211032)

<table>
<tr><th rowspan="2"></th><th colspan="3">Ancestral Generations</th></tr>
<tr><th>1</th><th>2</th><th>3</th></tr>
<tr><td rowspan="8">Sex ♀
No. 211032 Imp. Milk Maid</td><td rowspan="4">♂
No. 425404 Ireby Signet</td><td rowspan="2">No. 409267 ♂
Morning Sun</td><td>No. 409193 ♂
Inspector</td></tr>
<tr><td>No. —— ♀
Bessie 44th</td></tr>
<tr><td rowspan="2">No. —— ♀
Tulip 28th</td><td>No. 409093 ♂
Dainty Bean</td></tr>
<tr><td>No. —— ♀
Tulip 23d</td></tr>
<tr><td rowspan="4">♀
No. 211017 Border Lady</td><td rowspan="2">No. 433648 ♂
Border Stamp</td><td>No. 80356 ♂
Arkin Beau</td></tr>
<tr><td>No. —— ♀
White Sunshine</td></tr>
<tr><td rowspan="2">No. —— ♀
Lady Balmoral</td><td>No. 425402 ♂
Balmoral Pearl</td></tr>
<tr><td>No. —— ♀
Lady Benedict's Farewell</td></tr>
</table>

We see that in these two pedigrees there is, in the first ancestral generation, one ancestor (Ireby Signet) which occurs in both. Hence we have

$$K_1 = \frac{100\,(1)}{2} = 50.00.$$

In the second generation there are three ancestors (Morning Sun, Tulip 28th and Border Stamp), which occur in both pedigrees, whence it follows that

$$K_2 = \frac{100\,(3)}{4} = 75.00.$$

In the third generation there are six common ancestors

TABLE VII

PEDIGREE OF IMP. WHITE QUEEN (545726)

	Ancestral Generation		
	1	2	3
Sex ♀ Imp. White Queen No. 545726	♂ No. 425404 Ireby Signet	No. 409267 ♂ Morning Sun	No. 409193 ♂ Inspector
			No. ——— ♀ Bessie 44th
		No. ——— ♀ Tulip 28th	No. 409093 ♂ Dainty Bean
			No. ——— ♀ Tulip 23d
	♀ No. 501778 Border Queen	No. 433648 ♂ Border Stamp	No. 80356 ♂ Arkin Beau
			No. ——— ♀ White Sunshine
		No. ——— ♀ Diamond Queen	No. 501767 ♂ Levens Guardsman
			No. ——— ♀ Landford Diamond

(all involved from the second ancestral generation) and hence

$$K_3 = \frac{100\,(6)}{8} = 75.00.$$

So that we may say that Imp. Milk Maid and Imp. White Queen are 50 per cent. related in the first ancestral generation, and 75 per cent. in the second and third. This case will illustrate the superiority of the present exact numerical expression of relationship over the ordinary verbal expression. These two cows are half sisters, both having the same sire (this degree of relationship is indicated numerically always by $K_1 = 50$). But they are more closely related than two individuals which are *only* half sisters, because they have *also* one grandsire (Border Stamp) in common. Their total degree of relationship is simply not expressible verbally, by any term of kinship known to me in the English language. Yet by the method

here described it is exactly expressible in the form $K_1 = 50$, $K_2 = K_3 = 75$.

5. It will be perceived that the form of relationship coefficient here proposed leads to precisely the same numerical results in simple pedigrees, with not too involved inbreeding or kinship, as that given in my former paper[7] except for the fact that I have here changed the subscript designation of the K's to bring them into conformity with the total inbreeding coefficients. The earlier form proposed for these coefficients would always give the same numerical values as the present one if certain rather complicated rules of application, which were not clearly or rigorously set forth in the earlier paper, were to be followed. But the present simplified form does away entirely with the need for these complicated rules of procedure.

6. It is of interest to set forth in tabular form the values of the relationship coefficients for the commonly recognized degrees of kinship. This is done in Table VIII, in which the different degrees of kinship are arranged in descending order of closeness, in general. In some cases, as, for example, parent and offspring and half brothers (or half sisters), groups of two or three different sorts of kinship showing the same numerical degree of relationship should be regarded as bracketed, since there is no more reason for placing one of these first than another.

From this table a number of interesting points emerge. We note that the absolute maximum of closeness of relationship is that of brother and sister. The parent and offspring relationship is one half as close. Uncle and nephew (or niece), or single first cousins, are twice as closely related as grandparent and offspring. Some of these comparisons made obvious by the table may seem at first thought to give unexpected results, but if one will take the trouble to write down pedigrees for the stated

[7] Pearl, R., AMER. NAT., Vol. XLVIII, pp. 513–523, 1914.

degree of kinship, he will see upon careful consideration the reasonableness of the numerical result.

TABLE VIII

VALUES OF THE RELATIONSHIP COEFFICIENTS FOR VARIOUS DEGREES OF KINSHIP

Degree of Kinship	K_1	K_2	K_3	K_4
Brother and brother (or sister)	100	100	100	100
Parent and offspring	50	50	50	50
Half-brother and half-brother (or half-sister)	50	50	50	50
Double first cousins	0	100	100	100
Single first cousins	0	50	50	50
Uncle and nephew (or niece)	0	50	50	50
Grandparent and offspring	0	25	25	25
Quadruple second cousins	0	0	100	100
Double second cousins	0	0	50	50
Single second cousins	0	0	25	25
Propositus and first cousin once removed	0	0	25	25
Propositus and first cousin twice removed	0	0	0	12.5

7. There are two points in the development of relationship coefficients in this paper which may seem open to criticism. The first is that according to the definitions and formulæ of this paper, the degree of relationship between two individuals is not affected by the number of times *the same* common ancestor occurs in the pedigree of either of the two individuals. The fact that such ancestor occurs at least *once* in both pedigrees makes it a common ancestor. If it occurred more times it would not be a *more* common ancestor, because after all it would still be, all the time, just the same identical individual, made up of the same germ plasm. Put in another way, it is *community* of ancestry of two individuals which makes kinship. But the multiple appearance of the *same* individual in two pedigrees does not make any more ancestors common to the two related individuals than if this ancestor occurred only once in each pedigree. Consider an individual *A* which is rather intensely inbred with reference to an ancestor *X*. Consider another individual *B* which is also inbred to some extent with reference to the same individual *X*. Because they have a common ancestor *X*, *A* and *B* are related. But, according to the concep-

tion on which the present method of measuring kinship is based, the fact that A and B happen both to be inbred in respect to X, does not make them any more closely related to each other than if they were not so inbred. It may be of interest in this connection to point out, not as adding to the scientific exactitude of the position here taken, but as indicating what the common sense of men who have given thought to the subject of consanguinity has been, that the position here adopted that in determining degree of kinship a common ancestor counts but once as such, appears to be exactly in agreement with the position of both the canon law and the civil law on the same point.

The second point in regard to which criticism might seem to be possible is the method of referring the inbreeding or relationship to the ancestral generations. In all of these Studies the inbreeding or relationship is referred to the generation of the more remote (from the propositus) of the two appearances in a pedigree of a repeated ancestor. The logic of this procedure, rather than the reverse, is found in the circumstance that the fact of inbreeding (or kinship) does not establish itself until the more remote reappearance is reached. Thus it is impossible to know that a mating is of uncle and niece until the grandparental generation is reached.

$$
\begin{array}{l}
a\begin{cases} b \\ c \end{cases} \\
x\begin{cases} d\begin{cases} b \\ c \end{cases} \\ e \end{cases}
\end{array}
$$

Ancestral generation . . . 1 2

a is the uncle of x, the common ancestors being b and c, but this fact is not known until the second ancestral generation is reached. The only logical method of representing these facts exactly in a numerical way would seem to be to say, in effect, that up to and including the first ancestral generation of a and x there is no evidence that these individuals are at all related, and therefore

$K_1 = 0$. In the second ancestral generation, on the contrary, it appears that two ancestors, b and c, in the pedigree of x are the same individuals as appeared in the first ancestral generation of a. Therefore it now appears that a and x are related to the extent of 50 per cent. by the existence of community of ancestry in the second ancestral generation. It would seem only logical to attach the numerical measure of relationship to the generation in which it is first proved to exist. Again, this is precisely the point of view regarding the matter which has been taken by the canon law and Roman civil law.

These two points, which seem so obvious to the writer as to be difficult to discuss, are taken up here because correspondence has shown that they have been a source of difficulty with some who have undertaken the study of inbreeding in domestic animals by the methods set forth in these studies. It is hoped that the simpler and more precise definitions of both inbreeding and relationship constants given in this paper may help to clear up such difficulties, which must arise, it would seem, from a lack of a thorough grasp of the characteristics of pedigrees.

Summary

In this paper the basic concepts of inbreeding are redefined in a simple and rigorous manner, and on the basis of these definitions a new and more accurate method of measuring and expressing numerically the degree of kinship between any two individuals whatsoever, whose pedigrees are known, is set forth and illustrated by examples.

A new constant, the *partial inbreeding index*, is described. Its purpose is to indicate numerically the part of the total inbreeding exhibited in the pedigree of any individual which is due to relationship between the sire and the dam of that individual.